Conor R. Walker

New York Genome Center, 101 6th Ave, New York, NY 10013, USA

✉ conor@conorwalker.net

ⓘ 0000-0001-5617-5086 · ⓖ Conor R. Walker · ⓞ conorwalker · ⓛ conor-walker · 🐦 @BioConorWalker

Education

University of Cambridge, UK *2017–2021*
PhD Computational Biology
- Thesis: Statistical analysis of short template switch mutations in human genomes
- Advisers: Nick Goldman, Aylwyn Scally, Nicola De Maio

Newcastle University, UK *2016–2017*
MSc Bioinformatics · Distinction (83%)
- Thesis: Optimising nucleic acid sequences for DNA strand displacement systems
- Adviser: Harold Fellermann

Liverpool John Moores University, UK *2013–2016*
BSc (Hons) Zoology · First-class honours (80%)
- Thesis: The genetics of insecticide resistance in the blackfly *Simulium vittatum*
- Advisers: Craig Wilding, Will Swaney

Research experience

New York Genome Center, USA *Jan. 2022–Present*
Postdoctoral Research Associate
- In the G^2 Lab, I am leading research into privacy-preserving methods for analyzing large-scale human functional genomics data. Joint appointment at Columbia University

Columbia University - Department of Biomedical Informatics, USA *Jan. 2022–Present*
Postdoctoral Research Fellow

European Bioinformatics Institute (EMBL-EBI), UK *2017–2022*
Predoctoral Researcher
- I developed statistical methods to identify template switch variants in DNA sequences
- I characterised the human evolutionary, population, and cancer landscape of template switch mutagenesis
- I designed convolutional neural networks for accurately detecting between-species positive selection, validated through large-scale simulations
- I co-led several projects characterizing global SARS-CoV-2 sequence evolution

This work was funded by EMBL and NIHR BRC.

Newcastle University, UK *2016–2017*
Postgraduate Student Research
- I implemented multi-objective genetic algorithms for designing and optimising DNA sequences for use in DNA computing systems

Liverpool John Moores University, UK *2016*
Undergraduate Student Research
- I identified enzyme families and point mutations associated with insecticide resistance in a medically-important blackfly genome using phylogenetic methods

Publications

phastSim: Efficient simulation of sequence evolution for pandemic-scale datasets.
De Maio N., Boulton W., Weilguny L., **Walker C. R.**, Turakhia Y., Corbett-Detig R., Goldman N.
PLOS Computational Biology, 18, e1010056 (2022).

Mutation rates and selection on synonymous mutations in SARS-CoV-2.
De Maio N., **Walker C. R.**, Turakhia Y., Lanfear R., Corbett-Detig R., Goldman N.
Genome Biology and Evolution 13, evab087 (2021).

Short-range template switching in great ape genomes explored using pair hidden Markov models.
Walker C. R., Scally A., De Maio N., Goldman N.
PLOS Genetics 17, e1009221 (2021).

Stability of SARS-CoV-2 phylogenies.
Turakhia Y., De Maio N., Thornlow B., Gozashti L., Lanfear R., **Walker C. R.**, Hinrichs A. S., Fernandes J. D., Borges R., Slodkowicz G., Weilguny L., Haussler D., Goldman N., Corbett-Detig R.
PLOS Genetics 16, e1009175 (2020).

A phylodynamic workflow to rapidly gain insights into the dispersal history and dynamics of SARS-CoV-2 lineages.
Dellicour S., Durkin K., Hong S. L., Vanmechelen B., Martí-Carreras J., Gill M., Meex C., Bontems S., André E., Gilbert M., **Walker C. R.**, De Maio N., Faria N., Hadfield J., Hayette M., Bours V., Wawina-Bokalanga T., Artesi M., Baele G., Maes P.
Molecular Biology and Evolution 38, 1608–1613 (2020).

Masking strategies for SARS-CoV-2 alignments.
De Maio N., **Walker C. R.**, Borges R., Weilguny L., Slodkowicz G., Goldman N.
virological.org/t/masking-strategies-for-sars-cov-2-alignments/480 (2020).

Issues with SARS-CoV-2 sequencing data.
De Maio N., **Walker C. R.**, Borges R., Weilguny L., Slodkowicz G., Goldman N.
virological.org/t/issues-with-sars-cov-2-sequencing-data/473 (2020).

Selected presentations

RECOMB 2023 *Apr. 2023*
Poster: Private information leakage in single-cell omics data
Walker C. R., Gürsoy G.

Rocky 2022 *Dec. 2022*
Poster: Private information leakage in single-cell omics data
Walker C. R., Gürsoy G.

ISMB/ECCB 2021 *Jul. 2021*
Talk: Accurate detection of positive selection using convolutional neural networks
Walker C. R., De Maio N., Goldman N.

SMBE 2021 *Jul. 2021*
Talk: Accurate detection of positive selection using convolutional neural networks
Walker C. R., De Maio N., Goldman N.

SMBE 2019 *Jul. 2019*
Poster: Short template switches explain mutation clusters in the human genome
Walker C. R., De Maio N., Goldman N.

NIHR BRC Annual Research Day *Jul. 2018*
Talk: Investigating short template switch mutations in humans and model organisms
Walker C. R.

Teaching

Students supervised

Marvin Limpijankit · Undergraduate student *2022*
• Project: Predicting true somatic mutations using machine learning

Fatma Rabia Fidan · Masters student *2021*
- Project: Identifying *de novo* template switch mutations in trios of human genomes

Shayesteh Arasti · Masters student *2021*
- Project: Using convolutional neural networks to detect positive selection

Viacheslav Vasilev · Undergraduate student *2020*
- Project: Using convolutional neural networks to detect positive selection

William Xu · Undergraduate student *2020*
- Project: Using convolutional neural networks to detect positive selection

Teaching assistant

Medics to Coders · University of Cambridge, UK *2018*
- I delivered practical sessions on programming at the School of Clinical Medicine

Honours and awards

ISMB/ECCB 2021 EvolCompGen - Best Talk Award *Jul. 2021*

EMBL-EBI PhD Seminar Day - Best Talk Award *May 2019*

MSc Bioinformatics Prize - highest overall grade on my master's degree program *Oct. 2017*

BSc Zoology Prize - highest overall grade on my bachelors's degree program *Sep. 2016*

Other scientific activities

Committees

Student Representative · EMBL-EBI, UK *2018–2020*
- Attended meetings with senior management from across all EMBL outstations to discuss issues impacting the graduate community at EMBL
- Organised a variety of academic and social events within the institution

Graduate Student and Post-Doc Forum member · University of Cambridge, UK *2018*
- I worked as part of a group of students and postdocs to devise strategies for creating a more positive experience for graduate students in the School of Life Sciences

Memberships

International Society for Computational Biology
Society for Molecular Biology and Evolution
Society for the Study of Evolution

Reviewer for

Nature Medicine
Molecular Biology and Evolution
BMC Bioinformatics

Skills

Programming	Python, Bash, C++, R, AWK, LaTeX
Libraries	Numpy, SciPy, Keras, TensorFlow, PyTorch, pandas, matplotlib, seaborn, Dask, Numba, pysam, scikit-allel, Scanpy, scvi-tools
Coding practices	Git, Snakemake, Nextflow, unit testing, Docker, Singularity, Jupyter, Vim
High performance computing	LSF, Slurm
Cloud computing	AWS, Google Cloud
Operating systems	Linux, MacOS